DEPARTMENT OF STATISTICAL SCIENCES, UNIVERSITY OF TORONTO

Analysis on Excessive Deaths in Quebec During COVID-19 Based on Time Series with Bayesian Inference

Jiaqi Bi
February 17, 2022

ABSTRACT The COVID-19 is the most horrified disaster in recent years, where lots of people's living style, economy, and health are at risks. The report studies about the deaths caused by COVID-19, and the excessive deaths without COVID-19 impacts, as well as the vaccination efficacy on reducing the death rates in Quebec province. The report uses the method of Time Series based on Bayesian inference. The result shows that elderly people are more susceptible on severe health conditions caused by COVID-19. The vaccination helps prevent the death rate to increase as shown in the result as well.

1 INTRODUCTION

As the global pandemic keeps affecting our daily lives, and hospitals are running overloaded, a concern of the death rate on the COVID-19 has raised. The report will analyze the excessive deaths in Quebec province of Canada using semi-parametric model based on Bayesian inference to predict how the death rates are being affected comparing to non-COVID eras. The public health sectors proposed the pandemic is more hazardous for middle-aged and elderly people, because those who often have chronic disease such as high blood pressure, coronary heart disease or weaker constitutions.

2 METHOD

2.1 DATA VISUALIZATION

The death statistics seem to have some cycled and time-related factors as shown in the total deaths at Figure 2.1. Moreover, elderly (shown in the color of light blue) and middle aged groups (shown in the color of green) have higher deaths than people under 49 years old (shown in the color of red). Especially the elderly group, that has more than 50% contribution on the total deaths (shown in the color of purple).
The death counts have an overall increasing trend from 2010 to 2022, but it has seasonal effects that some may believe it is due to seasonal disease such as influenza. However, the total death counts reached 2000 which is the highest point in 2020 as shown in Figure 2.2. The weekly death data in Figure 2.3 shows a tremendous high rates of death in the plot of 2020 that it does not seem to follow the usual death trend.

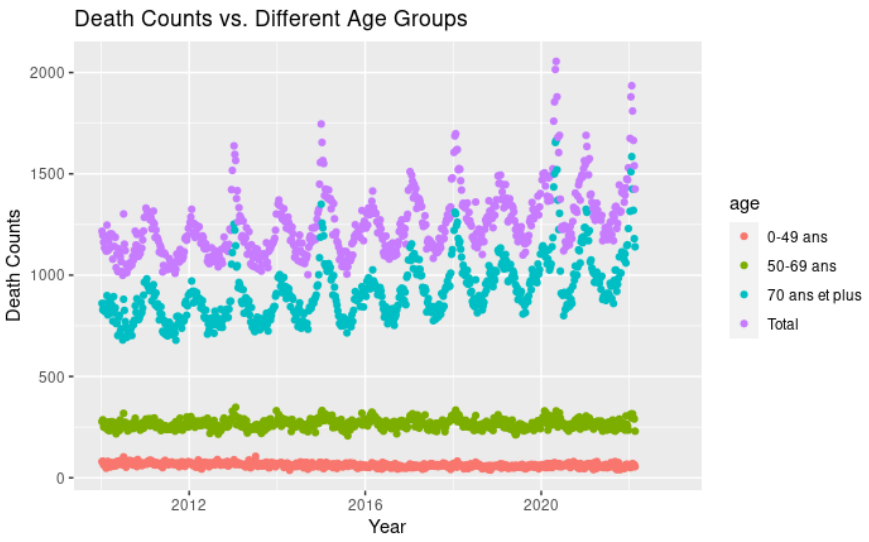

Figure 2.1: Death Counts with Different Age Groups

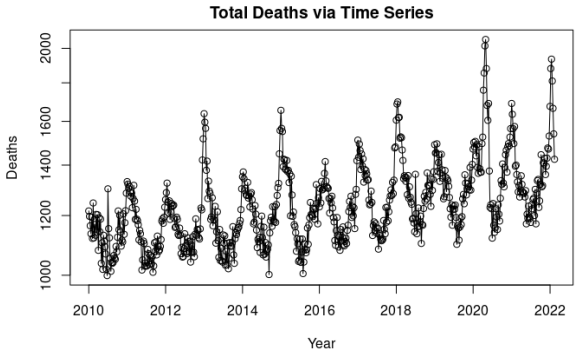

Figure 2.2: Total Deaths

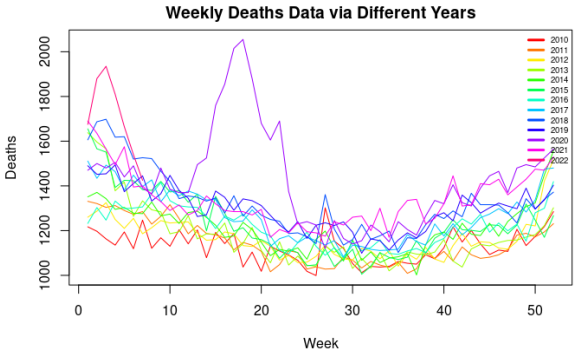

Figure 2.3: Weekly Total Deaths

2.2 Excessive Deaths

Now, using the average of the pre-2020 death statistics excluding 2020 and using the result to get subtracted by the 2020 death data may be sufficient to calculate the excessive deaths:

$$\mathbf{D}_{2020} - \frac{\sum_{i=2010}^{2019} \mathbf{D}_i}{(2019-2010)}$$

where $\mathbf{D}$ denotes the data matrix of year i.
Based on the result of the excessive deaths, we discover the excessive deaths reached over 800 counts and it is much higher than others. However, the elderly people are much more susceptible in Figure 2.4 where other age groups do not attempt to be highly affected. Thus, we say COVID-19 caused a lot of deaths for elderly people because the pandemic starts from the beginning of 2020. The trend of deaths keep raising with time due to the aging population keeps increasing in recent years.

2.3 Statistical Analysis

In order to make a comprehensive data analysis, we are going to use smoothing time series model prior to forecasting. Due to the seasonal effects shown in the previous graphs, we are going to analyze the sinusoidal trends in the graph. The sinusoidal function is defined as:

$$s(t) = a\cos(2\pi t/365.25 + \phi)$$

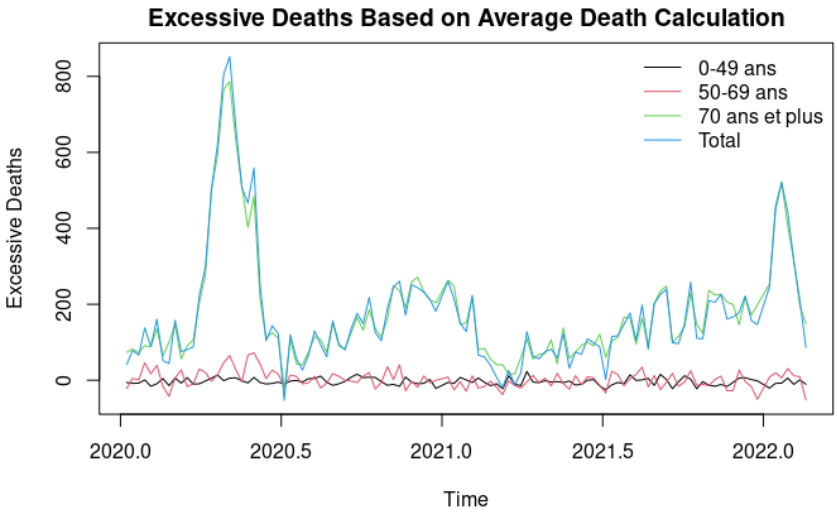

Figure 2.4: Excessive Deaths Based on Average Calculation

such that a is the amplitude, and ϕ is the phase. The frequency for this setting is $1/365.25$, and the cycle completes every 12 months. We can use the following proposed equation to deal with the seasonal effect:

$$\beta_1 \cos(2\pi t/365.25) + \beta_2 \sin(2\pi t/365.25) = \sqrt{\beta_1^2 + \beta_2^2} \cos(2\pi t/365.25 + \arctan(-\beta_2/\beta_1)$$

Furthermore, we are going to create some time covariates within the time series model:

$$X_{i1} = \sin(2\pi t_i/365.25)$$
$$X_{i2} = \sin(4\pi t_i/365.25)$$
$$X_{i3} = \cos(2\pi t_i/365.25)$$
$$X_{i4} = \cos(4\pi t_i/365.25)$$

where t denotes the certain time. Note that X_{i1} and X_{i3} are a cycle with 12 months, and the other two variables denote the cycle of 6 months. Now we can construct the semi-parametric time series model with Bayesian inference. We assume

$$Y_i \sim \text{Poisson}(\lambda_i)$$

and

$$\log(\lambda_i) = X_i\beta + U_i + V(t_i)$$

where λ_i is the expected value at certain time t_i. β is the parameter with priors we propose as

$$\beta \sim N(0.01, 0.5)$$

Moreover, we have random effects including time random effects $V(t_i)$ and normal random effects U_i such that

$$U_i \sim N(\log(1.2), \sigma_U^2)$$

and

$$[V_1 ... V_T]^T \sim \text{RW2}(0, \sigma_V^2)$$

where those random effects have priors being set as

$$P(\sigma_U > \log(1.2)) = 0.5$$

and

$$P(\sigma_V > 0.01) = 0.5$$

With all above parameters and models, we are feasible to forecast the 2021 and 2022 death data without counting COVID-19 effect into it.

3 Results

Based on Table 3.1, we discover that the parameter is relatively high which indicates the death counts will increase with years. The time random effects play a role that proves the assumption of increasing trend. The 80% credible interval for the parameter and time random effects are statistically significant.

Table 3.1: Parameters Posterior Quantile Statistics

	0.5 Quantile	0.25 Quantile	0.975 Quantile
Parameter	7.1029	7.0611	7.1424
X_{i1}	0.0513	0.0448	0.0578
X_{i2}	0.0121	0.0064	0.0177
X_{i3}	0.0972	0.0907	0.1038
X_{i4}	0.0131	0.0074	0.0187
SD Random Effects	0.0358	0.0322	0.0397
SD Time Random Effects	0.1372	0.0889	0.2102

Figure 3.1 shows a COVID-19 vaccination status in Quebec, as we can observe that with the increase of the vaccined population, the death counts drop in Figure 3.2 where the start date of vaccination is marked as a vertical green line. Especially the elderly group of people have less risks on COVID-19 than before due to vaccinations.

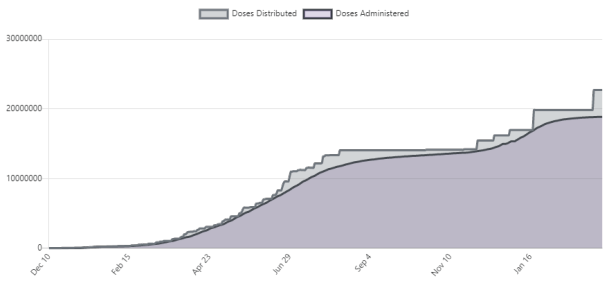

Figure 3.1: Quebec Vaccine Status

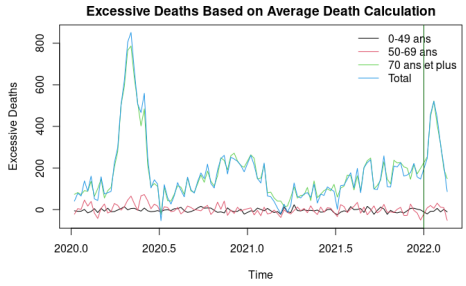

Figure 3.2: Deaths after Vaccines

The forecasting of the death data without the effect of COVID-19 is shown in Figure 3.3 where there does not seem to be a much higher deaths in 2020. The death counts continue to raise with time. The increasing rate has seasonal effects as shown in Figure 3.4 where there is not a high increasing rate comparing to previous years for 2020. Therefore, this proves that COVID-19 has huge impacts on total death population, especially elderly groups.

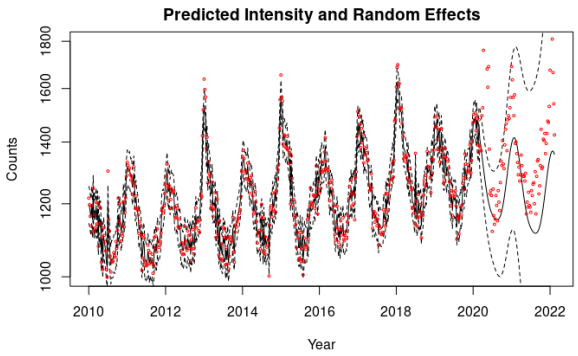

Figure 3.3: Predicted Plot

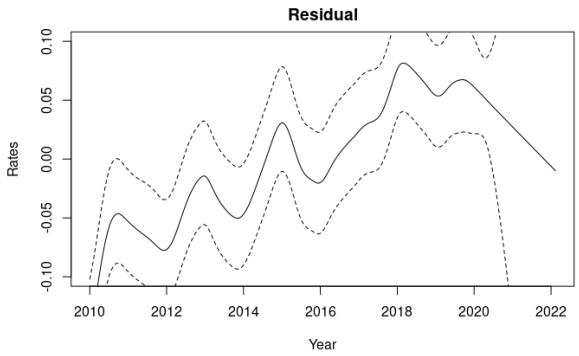

Figure 3.4: Residual

Figure 3.5 and Figure 3.6 show a comparison between real data and predicted data where the difference is caused by COVID-19. The red dots are based on the real data, that shows a much higher counts from 2020 to 2022. Looking closely to the data of 2020 in Figure 3.6, the death population rapidly increased during the beginning of 2020.

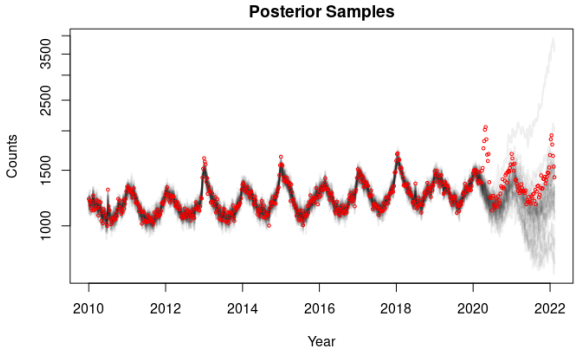

Figure 3.5: Posterior Samples 1

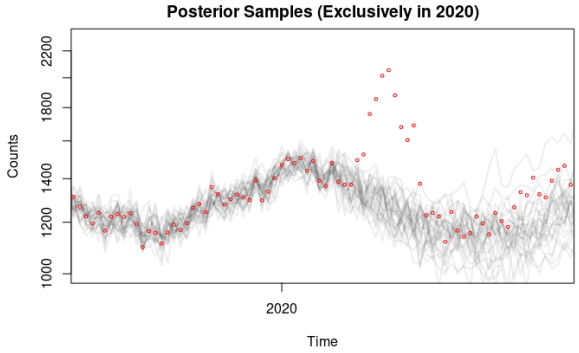

Figure 3.6: Posterior Samples 2

As shown in Figure 3.7 and Figure 3.8, the COVID-19 hits the highest death counts in 2020. Afterwards, the death counts are within 0 to 500 in Figure 3.8.

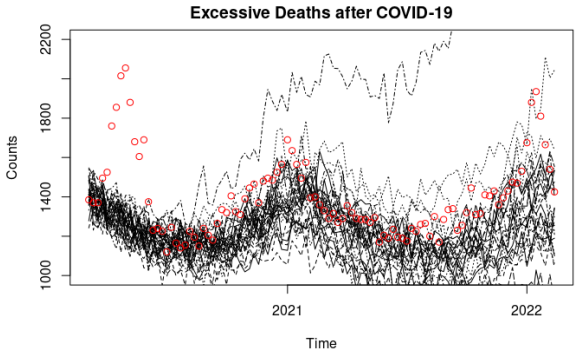

Figure 3.7: Trend after COVID-19

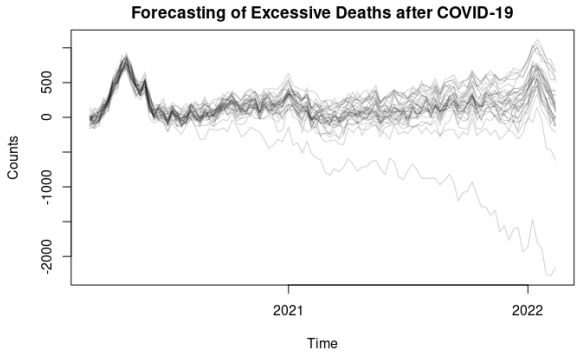

Figure 3.8: Forecast after COVID-19

4 Conclusion

In conclusion, the COVID-19 indeed affects the death counts especially in 2020 for elderly people. Meanwhile, the vaccination rates increase in 2021 and 2022, this leads to a lower increasing rate of

deaths caused by COVID-19. According to the forecasting data, we found that COVID-19 has a huge impact on the total deaths that the death counts in 2020 for real Quebec death data are much higher than the predicted data in 2020 for the model we proposed without COVID-19. Thus, we conclude that elderly people are more susceptible and more risky on the death caused by COVID-19, and we can conclude that the vaccination has huge positive effects on prevention of severe COVID-19 for susceptible group.

5 Limitations and Discussions

Due to the lack of information access on cause specific deaths, we are unable to analyze some other infectious, or serious illness may have caused a large deaths in 2020 such as influenza or cancers. Moreover, the vaccination does not represent the whole elderly group since there are more people have vaccination hesitancy. Therefore, there can be some future research on COVID-19 vaccine hesitancy based on their reasons of hesitancy.